# Optimal Sampling Strategies for Oceanic Applications

Dr Peter R. Oke
CSIRO Marine and Atmospheric Research and Wealth from Oceans Flagship Program
GPO Box 1538 Hobart TAS 7001 Australia
phone: (+61 3) 6232-5387 fax: (+61 3) 6232-5123 email: peter.oke@csiro.au

Award Number: N000014-07-11054,
including partial support from ONR Global's Naval International
Cooperative Opportunities Program (NICOP)
http://www.cmar.csiro.au/staff/oke/onr.htm

## LONG-TERM GOALS

The long-term goals of this project are to improve our ability to monitor and predict the ocean circulation through the assessment and design of ocean observing systems; and through the development of practical methods for ocean data assimilation.

## OBJECTIVES

We seek to develop new approaches, and to apply existing methods, to the problem of observing system design and assessment. This involves the establishment of a methodology and practical capability to support the routine assessment of the ocean observing system; to apply these capabilities to gain insight into the potential benefits of future observation types for data assimilating models; and to support operational groups in the design of new observing systems, either for specific targeted activities or for monitoring programs.

## APPROACH

A range of different approaches to observing system design and assessment have been exploited under this project. These include traditional methods, including Observing System Experiments (OSEs) and Observing System Simulation Experiments (OSSEs), and new methods that draw on aspects of data assimilation theory. We have applied these tools to the assessment and design of ocean observing systems on a range of scales and applications. This includes an assessment of the relative importance of different components of the global ocean observing system (GOOS), the design of a tropical mooring array, and an assessment of the potential benefits of different shelf observation platforms to a data assimilating model. A summary of OSE and OSSE activities conducted under the Global Ocean Data Assimilation Experiment (GODAE), an international effort in the development of ocean forecast and analysis capabilities, is described by Oke et al. (2009a).

Several aspects of ensemble data assimilation have been explored under this project. This includes an effort to understand the impact of using different formulations of ensemble square root filters, as well

# Report Documentation Page

*Form Approved*
*OMB No. 0704-0188*



| | | |
|---|---|---|
| 1. REPORT DATE<br>**2009** | 2. REPORT TYPE | 3. DATES COVERED<br>**00-00-2009 to 00-00-2009** |

| | |
|---|---|
| 4. TITLE AND SUBTITLE<br>**Optimal Sampling Strategies for Oceanic Applications** | 5a. CONTRACT NUMBER<br>5b. GRANT NUMBER<br>5c. PROGRAM ELEMENT NUMBER |
| 6. AUTHOR(S) | 5d. PROJECT NUMBER<br>5e. TASK NUMBER<br>5f. WORK UNIT NUMBER |
| 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)<br>**CSIRO Marine and Atmospheric Research and Wealth from Oceans Flagship Program,GPO Box 1538 Hobart TAS 7001 Australia,** | 8. PERFORMING ORGANIZATION REPORT NUMBER |
| 9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) | 10. SPONSOR/MONITOR'S ACRONYM(S)<br>11. SPONSOR/MONITOR'S REPORT NUMBER(S) |

12. DISTRIBUTION/AVAILABILITY STATEMENT
**Approved for public release; distribution unlimited**

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT | 18. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON |
|---|---|---|---|---|---|
| a. REPORT<br>**unclassified** | b. ABSTRACT<br>**unclassified** | c. THIS PAGE<br>**unclassified** | **Same as Report (SAR)** | **11** | |

**Standard Form 298 (Rev. 8-98)**
Prescribed by ANSI Std Z39-18

as the development of a new EnKF algorithm. These activities have been conducted with the aid of a series of small models.

Dr. Peter Oke is the P.I. on this project and leads the ocean data assimilation activities at the Centre for Australian Climate and Weather Research (CAWCR), a partnership between CSIRO and the BoM. Other researchers contributing to this research include Dr. Pavel Sakov, now based at NERSC in Norway, Dr. Terence O'Kane from CAWCR, Dr. Hans Ngodock and Dr. Gregg Jacobs from NRL.

**WORK COMPLETED**

*Observing System Design and Assessment*

We have exploited both OSEs and OSSEs under this project. OSEs generally involve the systematic denial, or withholding, of different observation types from a data assimilating model in order to assess the degradation in quality of a forecast or analysis when that observation type is not used. OSEs are most suited to quantifying the benefits of existing observation types for data assimilation models. Examples of OSEs performed under this project are documented by Oke and Schiller (2007a).

OSSEs, on the other hand, typically involve some sort of twin experiment where synthetic observations, usually extracted from a model, are assimilated into an alternative model. OSSEs are therefore well suited to evaluating the potential benefits of hypothetical observations that may not exist yet. Thus, these methods can be used to contribute to the design of future observing systems, quantifying their possible impacts and limitations. Examples of OSSEs performed under this project are documented by Oke et al. (2007b). These examples contributed to the design of the tropical Indian Ocean mooring array.

A new capability has been developed under this project that is built on the work of Bishop et al. (2001). The method seeks to identify the set of observations that minimise the analysis error variance for a predefined variable or quantity, given an ensemble of perturbations that represent the system's background error covariance. The approach is very general and can be applied to either a stationary or time-varying ensemble. Applications using a stationary ensemble are most applicable to models using a 3d assimilation system, like 3d-Var or EnOI. Applications to a time-varying system are most applicable to models using a 4d assimilation system, like 4d-Var or the ensemble Kalman Filter (EnKF). This method is documented by Sakov and Oke (2008a) and has been used to contribute to the design and assessment of the New South Wales (NSW) node of the Australian Integrated Marine Observing system (IMOS). The application to NSW IMOS is documented by Oke et al. 2009b).

*Data Assimilation*

Under this project, a fundamental exploration of ensemble data assimilation has also been undertaken. The significance of this work is best understood in light of the history of ensemble data assimilation that is summarized below. The ensemble Kalman filter (EnKF) was first proposed by Evensen (1994). The idea of the EnKF was that an ensemble of model perturbations is integrated - implicitly representing the system's background error covariance. This covariance is used to assimilation observations. At each analysis step the ensemble is updated, implicitly representing the system's analysis error covariance. Conceptually, the framework for the EnKF has not changed since its conception. However, the formulation proposed by Evensen was flawed, and the ensemble typically collapsed owing to the under-representation of the ensemble's analysis error covariance. Subsequently, Burgers et al. (1998) proposed a different formulation of the EnKF where observations are treated as random variables, and are perturbed. In theory, Burgers' approach, known as the perturbed obs EnKF, should yield an ensemble with the correct analysis error variance. But this approach is statistical and therefore suffers from sampling error - particularly for small ensemble size. This renders the perturbed obs EnKF sub-optimal in practice. Ensemble square root filters (ESRFs; e.g., Tippett et al. 2003) then became popular in ocean and atmospheric data assimilation because they are deterministic. They don't suffer from sampling error and are optimal for models where the ensemble size exceeds the dimension of the system's sub-space. Most of the development of ensemble-based methods has been conducted on small models (e.g., Lorenz and Emanuel 1998). However, when applying these methods to realistic models, for example general circulation models, in practice all ensemble-based filters are rank-deficient. That is, the ensemble size is typically much less than the dimension of the model's sub-space. So there aren't enough degrees of freedom in the ensemble to adequately span the model's error sub-space. This problem is overcome by using covariance localisation (e.g., Houtekamer and Mitchell 1998). Localisation is traditionally applied in grid point space. This is readily applied to the perturbed obs EnKF, but cannot formally be applied to ESRFs. Instead localised ESRFs are often implemented by assimilating only a sub-set of observations, local to each grid point, and neglecting observations outside a certain search radius (e.g., Anderson 2003; Ott et al. 2004). This approach is practical, but is not faithful to the theory that underpins ESRFs. Under this project, we have developed a new filter that is both deterministic, like the ESRFs, and readily permits localisation in grid-point space, like the perturbed obs EnKF. This filter is called the Deterministic EnKF (DEnKF). The DEnKF combines the numerical effectiveness, simplicity and versatility of the perturbed obs EnKF with the optimal performance of the ESRFs. The DEnKF is documented by Sakov and Oke (2008b).

We have investigated the impacts of using different ensemble transformations in ESRFs. It is found that only a sub-set of ensemble filters should be used in practice, and that some formulations significantly degrade the filter performance. This work is documented by Sakov and Oke (2008c).

**RESULTS**

*Observing system Assessment*

A series of OSEs have been performed using the Bluelink ReANalysis (BRAN; Oke et al. 2008) system. The Bluelink ocean data assimilation system (BODAS; Oke et al. 2005; 2008) that underpins BRAN is based on Ensemble Optimal Interpolation (EnOI). EnOI is well suited to OSEs – it is multivariate, using observations of one type to update variables of all types, and readily assimilates observations of different types in a single step. An example of the multivariate nature of EnOI is presented in Figure 1 showing the sea-surface temperature (SST), sea-level-anomaly (SLA) and sub-surface temperature (T) increments when different observations are assimilated. For this example, a series of warm-core and cold-core eddies are evident in the SLA observations, but only a cold-core eddy is evident in the SST observations. This can occur when, for example, the warm-core eddies are capped by cold near-surface waters. As a result, we find that when only altimetry is assimilated (Figure 1a,d,g) the increments reflect both the warm-core and cold-core eddies, with a clear surface-expression of the eddies in the SST increments. By contrast, when only SST is assimilated (Figure 1b,e,h), the increments reflect only cold-core eddies, plus a general T decrease over most of the region shown. When both altimetry and SST are assimilated (Figure 1c, f, i), the increments reflect both the warm-core and cold-core eddies, as well as the surface T decrease. This example demonstrates the importance of different types of observation for assimilation.

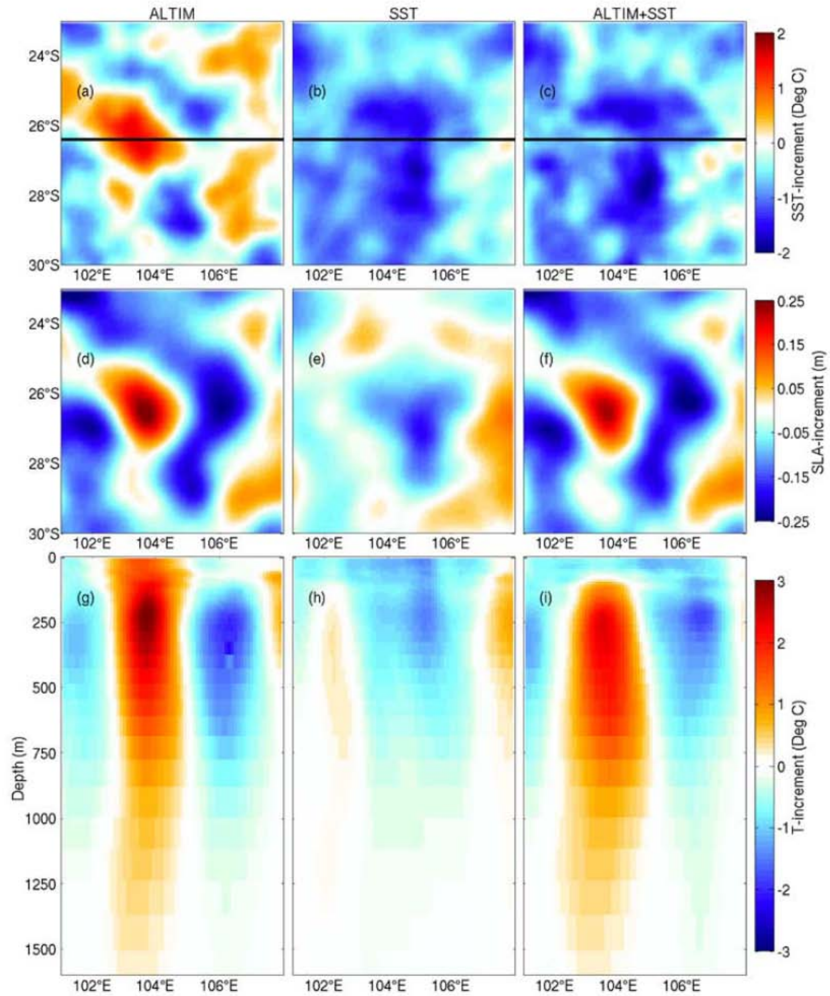


***Figure 1: Examples of (a, b, c) SST increments, (d, e, f) SLA increments, and (g, h, i) temperature increments for a longitude-depth section at 24.4°S (sections denoted in a–c) when only altimetry (a, d, g) only SST (b, e, h) and both altimetry and SST (c, f, i) observations are assimilated.***

***[Figure showing impact of different observation types]***

The OSEs presented by Oke and Schiller (2007) include 6 month data assimilating runs with different observation types assimilated, including: ALL (Argo+SST+ALTIM), NONE (no assimilation), Argo+SST, ALTIM+SST, and ALTIM+Argo. The RMS difference between the observed and reanalysed SLA and SST for each OSE is presented in Figure 2 and Figure 3 respectively. Reanalysed SLA fields are compared to along-track SLA (atSLA) from all available altimeters (Jason, Envisat and GFO). Reanalysed SST fields are compared to SST from AMSR-E. For comparison, the observed standard deviations are also shown in Figures 2 and 3 (for SST, this is based on anomalies from the seasonal cycle). The area-averaged RMS difference between observed and reanalysed sub-surface T and salinity (S) in the Australian region for each OSE is presented in Figure 4.

The OSE that assimilate all observations perform the best overall, with relatively small residuals for all variables. By contrast each OSE that assimilates only two of the three observation types performs poorly for at least one variable. For example, when only Argo and SST are assimilated, the residuals of SLA are significantly larger than ALL (Figure 2). Similarly, when only altimetry and SST are assimilated, the residuals in sub-surface T and S are large compared to ALL (Figure 4); and when only altimetry and Argo are assimilated, the residuals for SST become comparable, in some locations, to those in NONE (Figure 3).

Based on the OSEs presented by Oke and Schiller (2007), we conclude that each observation type brings complementary information to the GOOS. This result sends a clear message to policy makers that all observation types are needed to constrain mesoscale ocean models either for reanalysis or prediction.

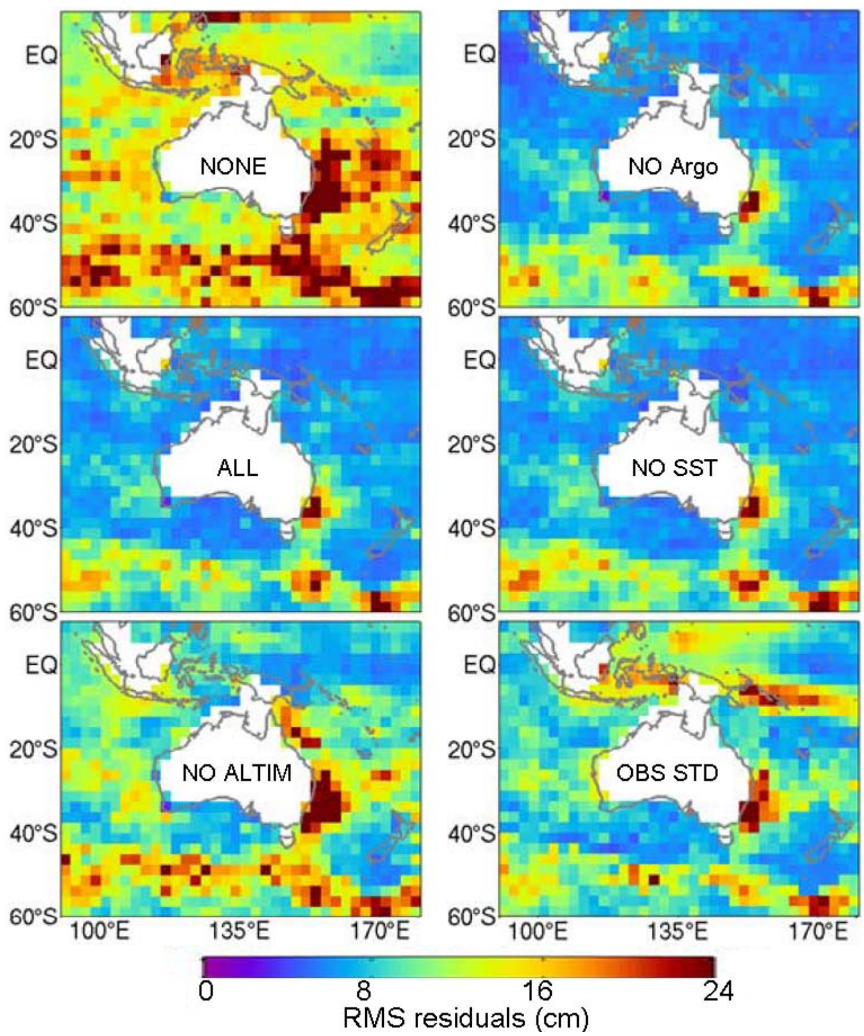


***Figure 2: RMS residuals between observed and modelled SLA for each OSE; and the observed standard deviation (bottom right). Statistics are computed using atSLA observations from Jason, Envisat, and GFO for the period January–May 2006.***
***[Figure showing relative SLA errors in an eddy-resolving model run when different observation types are assimilated]***

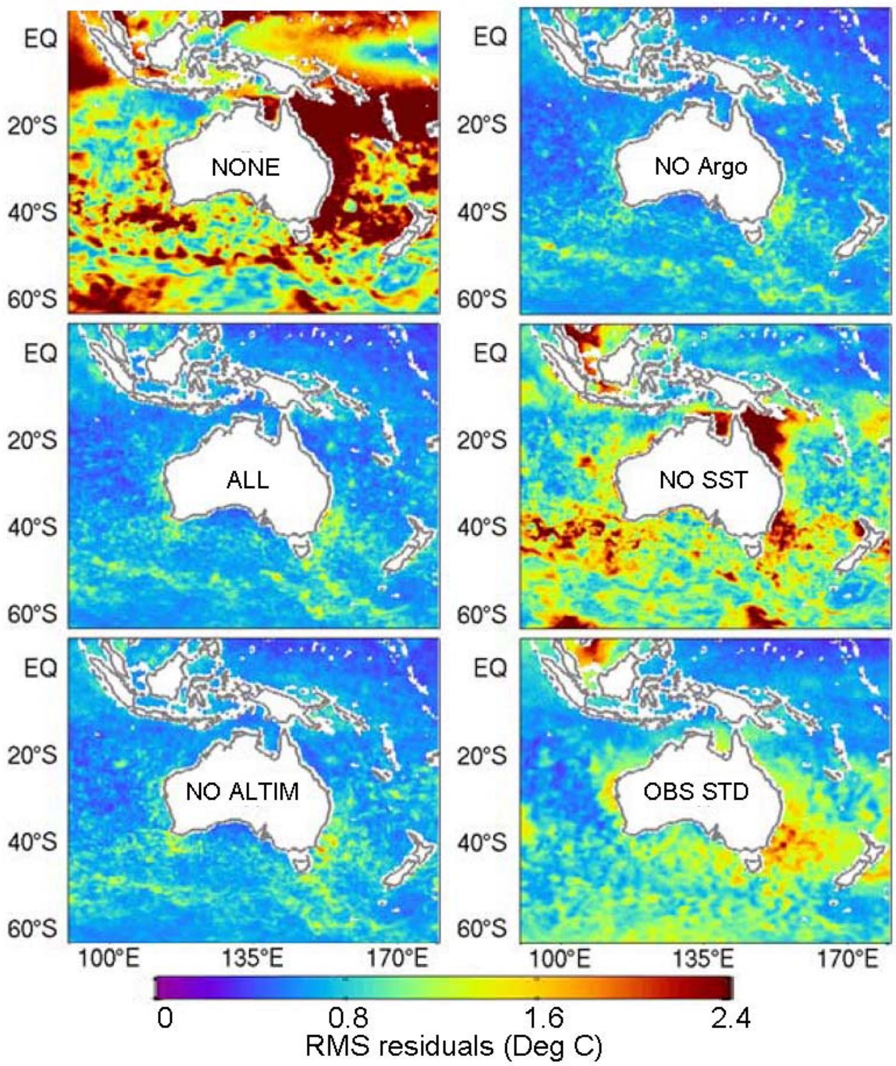


***Figure 3: RMS residuals between observed and modelled SST for each OSE; and the observed standard deviation (bottom right). Statistics are computed using AMSRE observations for the period January–May 2006.***
***[Figure showing relative SST errors in an eddy-resolving model run when different observation types are assimilated]***

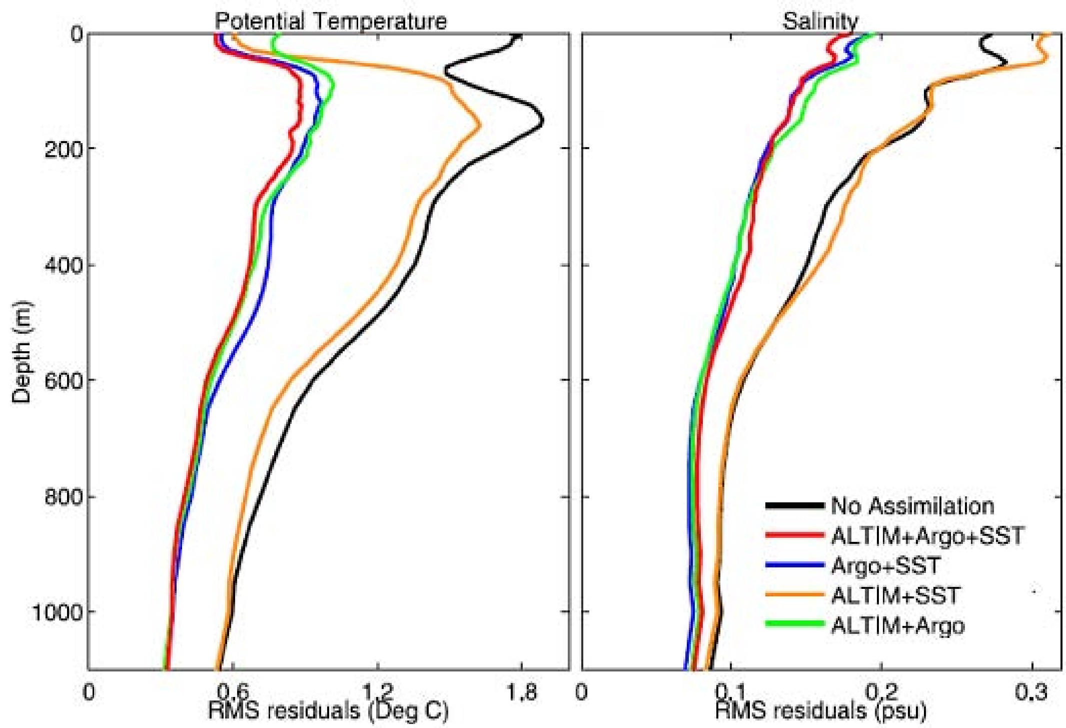


***Figure 4: RMS residuals between observed and modelled temperature (left) and salinity (right), for each OSE. Statistics are computed for the region 90–180°E and south of the equator using 3159 Argo profiles for the period January–May 2006.***
***[Figure showing relative temperature and salinity errors in an eddy-resolving model run when different observation types are assimilated]***

*Observing System Design*

The versatility of the ensemble-based method for observing system design developed here (Sakov and Oke 2008a) is demonstrated in the published literature through applications to the design of the tropical Indian Ocean (TIO; Sakov and Oke 2008a) and the NSW IMOS shelf observation program (Oke et al. 2009b).

The application to the TIO (Sakov and Oke 2008a) shows that the proposed TIO mooring array should produce results that compare well to the objectively designed arrays for intraseasonal variability. This study also identified regions in the east and northeast of the TIO and south of India as the most important regions for monitoring intraseasonal variability in the TIO. Observations near 9°S, where seasonal Rossby waves are prevalent, are also identified as key regions for monitoring seasonal-to-interannual variability.

The results of the application to NSW IMOS (Oke et al. 2009b) demonstrate that if high frequency (HF) radar observations are assimilated along with the standard components of the GOOS, the analysis errors are likely to reduce by as much as 80% for velocity and 60% for T, S and sea-level in the vicinity of the observations. This is an important result for Bluelink, since it impacts the planning of future Bluelink activities and helps set priorities for the uptake of new data types into the operational forecast system.

*Data Assimilation*

Sakov and Oke (2008b) present a number of experiments to compare the performance of the DEnKF with both the perturbed obs EnKF and an ESRF using three small models. It is shown that the DEnKF performs as well as the ESRF and is a significant improvement over the perturbed obs EnKF. It is also shown that under many circumstances, the DEnKF is more robust than ESRFs to mis-specification of tuning parameters, like inflation.

The paper by Sakov and Oke (2008c) considers implications of different forms of the ensemble transformation in the ensemble square root filters (ESRFs) for the performance of ESRF-based data assimilation systems. They show that it is important to use a mean-preserving solution for the ensemble transform matrix in ESRFs. Not doing so seriously degrades the performance of the ESRF.

## IMPACT/APPLICATIONS

The tools for objective array design that have been developed under this project are very powerful. Given a time series, or model ensemble, of oceanic fields for some region, these tools facilitate the efficient design and assessment of an observation array for that region. These calculations can be readily performed in real time, together with an EnKF, to undertake adaptive sampling. Our experiments indicate that the most critical aspect of any application is formulating the cost function. That is, in determining exactly what it is that we wish to monitor. For example, an “optimal” array for monitoring intraseasonal mixed layer depth is likely to be very different from an “optimal” array for monitoring interannual variability (Sakov and Oke 2008a). The results from the application of the ensemble-based method for observing system design to NSW IMOS have been used by the Bluelink science team to set priorities on the uptake on new observation types into the operational Bluelink forecast system.

The results from the OSEs presented by Oke and Schiller (2007) provide an important assessment of the performance of the GOOS for a data-assimilating, eddy-resolving ocean forecast system. Such information is essential for planning investment in the maintenance and development of the GOOS. The finding that different observation types provide complimentary information is an important result that should help persuade policy makers to maintain all of the observation programs.

Sakov and Oke (2008c) demonstrate that only mean-preserving ensemble transformations should be used by ESRFs. This is a very clear result, based on both theoretical and experimental evidence. This development has been adopted by the EnKF community (see http://enkf.nersc.no/Code/Analysis/meanpres.pdf). The development of the DEnKF provides a computationally efficient alternative to the ensemble data assimilation community. The uptake of the DEnKF has steadily increased since the publication of Sakov and Oke (2008b).

## RELATED PROJECTS

Bluelink is a partnership between CSIRO, the Bureau of Meteorology and the Royal Australian Navy. Many of the research activities undertaken in Bluelink have strong synergies for the project that is the subject of this annual report. The main objective of Bluelink is the development and application of an ocean forecast system for the mesoscale circulation around Australia. Applications of the Bluelink system are well documented (e.g., Oke et al. 2005; 2008; 2009c; Schiller et al. 2008). An example of results from the latest Bluelink reanalysis experiment is presented in Figure 5.

The Australian Integrated Marine Observing System (IMOS) program (www.imos.org.au). IMOS involves the provision of observational platforms (e.g., gliders, high-frequency radars, moorings) to establish a long-term monitoring capability for the oceans around Australia.

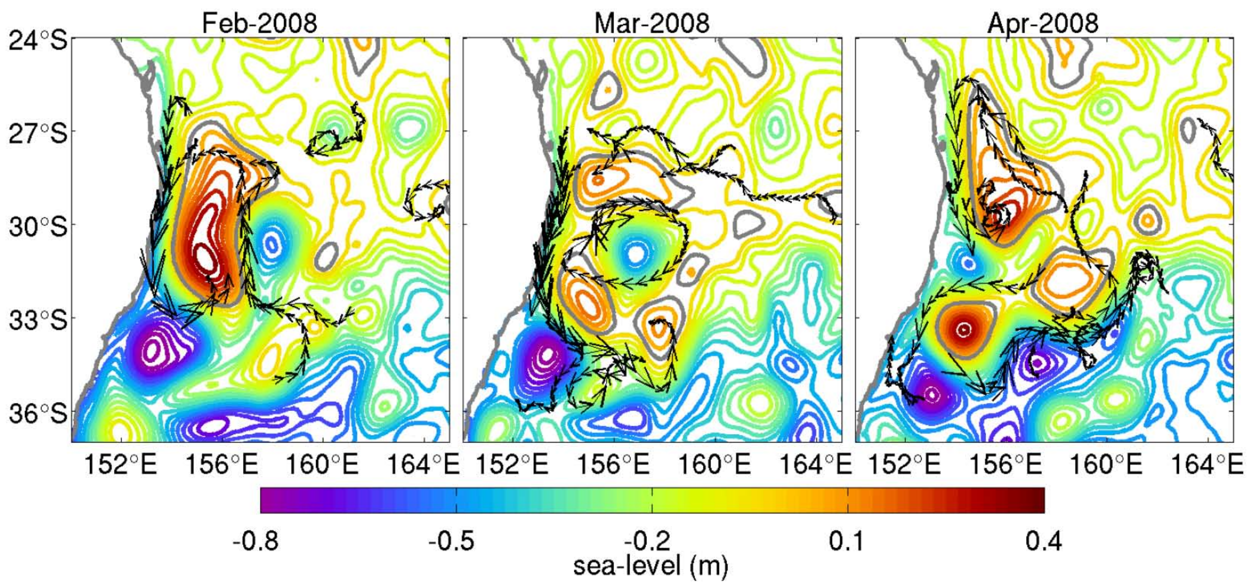


***Figure 5: Monthly mean sea-level from the latest Bluelink Reanalysis, with surface drifter velocities and trajectories overlaid.***

***[Figure showing good agreement between reanalyzed SLA and independent surface drifter trajectories in the latest Bluelink Reanalysis]***

## REFERENCES


Anderson, J. L., 2003: A local least squares framework for ensemble filtering. Monthly Weather Review, 131, 634-642.

Bishop, C., B. Etherton, and S. J. Majumdar, 2001: Adaptive sampling with the ensemble transform Kalman filter. part 1: theoretical aspects. *Monthly Weather Review*, **129**, 420–436.

Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Monthly Weather Review*, **126**, 796–811.

Lorenz, E. N., and K. A. Emanuel, 1998: Optimal sites for supplementary weather observations: Simulation with a small model. *Journal of Atmospheric Sciences*, **55**, 399-414.

Oke, P. R., A. Schiller, D. A. Griffin and G. B. Brassington, 2005: Ensemble data assimilation for an eddy-resolving ocean model, Quarterly Journal of the Royal Meteorology Society, 131, 3301-3311.

Oke, P. R., and A. Schiller, 2007a: Impact of Argo, SST and altimeter data on an eddy-resolving ocean reanalysis. *Geophysical Research Letters*, **34**, L19601, doi:10.1029/2007GL031549.

Oke, P. R., and A. Schiller 2007b: A model-based assessment and design of an optimal mooring array for the Indian Ocean. *Journal of Climate*, **20**, 3269-3283.

Oke, P. R., G. B. Brassington, D. A. Griffin and A. Schiller, 2008: The Bluelink Ocean Data Assimilation System (BODAS). Ocean Modelling, 20, 46-70.

Oke, P. R., M. Balmaseda, M. Benkiran, J. A. Cummings, E. Dombrowsky, Y. Fujii, S. Guinehut, G. Larnicol, P.-Y. Le Traon, and M. J. Martin, 2009a: Observing System Evaluations using GODAE systems, *Oceanography*, **22**(3), 144-153

Oke, P. R., P. Sakov and E. Schulz, 2009b: A comparison of shelf observation platforms for assimilation into an eddy-resolving ocean model. Dynamics of Atmospheres and Oceans, 48, 121-142, doi:10.1016/j.dynatmoce.2009.04.002.

Oke, P. R., G. B. Brassington, D. A. Griffin and A. Schiller, 2009: Data assimilation in the Australian Bluelink system. *Mercator Newsletter*, **34**, 35-44.

Ott, E., B.R. Hunt, I. Szunyogh, A.V. Zimin, E.J. Kostelich, M. Corazza, E. Kalnay, D.J. Patil, and J A. Yorke, 2004: A local ensemble Kalman filter for atmospheric data assimilation. Tellus, 56, 415-428.

Sakov, P., and P. R. Oke 2008a: Objective array design: Application to the tropical Indian Ocean. *Journal of Atmospheric and Oceanic Technology*, **25**, 794-807.

Sakov, P. and P. R. Oke 2008b: A deterministic formulation of the ensemble Kalman filter: an alternative to ensemble square root filters. Tellus-A, 60A, 361-371.

Sakov, P., and P. R. Oke 2008c: Implications of the form of the ensemble transformations in the ensemble square root filters. Monthly Weather Review, 136, 1042-1053.

Schiller, A., P. R. Oke, G. B. Brassington, M. Entel, R. Fiedler, D. A. Griffin, and J. V. Mansbridge, 2008: Eddy-resolving ocean circulation in the Asian-Australian region inferred from an ocean reanalysis effort. *Progress in Oceanography*, **76**, 334-365.

Tippett, M. K., J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker, 2003: Ensemble square root filters. *Monthly Weather Review*, **131**, 1485–1490.

## PUBLICATIONS

Oke, P. R., G. B. Brassington, D. A. Griffin and A. Schiller, 2009: Ocean Data Assimilation: a case for ensemble optimal interpolation. Australian Meteorological and Oceanographic Journal, in press. [in press, refereed]

Oke, P. R., M. Balmaseda, M. Benkiran, J. A. Cummings, E. Dombrowsky, Y. Fujii, S. Guinehut, G. Larnicol, P.-Y. Le Traon, and M. J. Martin, 2009: Observing System Evaluations using GODAE systems, *Oceanography*, **22**(3), 144-153. [published, refereed]

Oke, P. R., P. Sakov and E. Schulz, 2009: A comparison of shelf observation platforms for assimilation into an eddy-resolving ocean model. *Dynamics of Atmospheres and Oceans,* **48**, 121-142, doi:10.1016/j.dynatmoce.2009.04.002. [published, refereed]

Oke, P. R., G. B. Brassington, D. A. Griffin and A. Schiller, 2008: The Bluelink Ocean Data Assimilation System (BODAS). *Ocean Modelling,* **20**, 46-70. [published, refereed]

Oke, P. R., and P. Sakov 2008: Representation error of oceanic observations for data assimilation. *Journal of Atmospheric and Oceanic Technology,* **25**, 1004-1017. [published, refereed]

Sakov, P., and P. R. Oke 2008: Objective array design: Application to the tropical Indian Ocean. *Journal of Atmospheric and Oceanic Technology*, **25**, 794-807. [published, refereed]

Sakov, P., and P. R. Oke 2008: Implications of the form of the ensemble transformations in the ensemble square root filters. *Monthly Weather Review,* **136**, 1042-1053. [published, refereed]

Sakov, P. and P. R. Oke 2008: A deterministic formulation of the ensemble Kalman filter: an alternative to ensemble square root filters. *Tellus-A*, **60A**, 361-371. [published, refereed]

Oke, P. R., and A. Schiller, 2007: Impact of Argo, SST and altimeter data on an eddy-resolving ocean reanalysis. *Geophysical Research Letters,* **34**, L19601, doi:10.1029/2007GL031549. [published, refereed]

Oke, P. R., and A. Schiller, 2007: A model-based assessment and design of a tropical Indian Ocean mooring array, *Journal of Climate*, **20**, 3269-3283. [published, refereed]

Preprints of submitted manuscripts are available from http://www.cmar.csiro.au/staff/oke/.

## HONORS/AWARDS/PRIZES

Dr Peter Oke received a *Julius Career Award* from CSIRO. The *Julius Career Award* is designed to enhance the careers of early to mid-career scientists; and is intended to contribute towards their professional development.

Together with the Bluelink team, Dr Peter Oke was a finalist for the Eureka Prize for Outstanding Science in Support of Defense or National Security. Presented annually by the Australian Museum, *Eureka prizes* reward excellence in the fields of research and innovation, science leadership, school science and science journalism and communication.